Learning Seismocardiogram Beat Denoising Without Clean Data

1st Mohammad Nikbakht
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA, USA

2nd David J Lin
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA, USA

3th Omer T Inan
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA, USA

***Abstract*—Noninvasive monitoring of cardiovascular health plays a crucial role in predicting risks and reducing mortality rates, especially in the context of trauma care. The seismocardiogram (SCG) in particular is a noninvasive signal that has been shown to monitor key health parameters related to blood volume loss estimation, suggesting its ability to guide trauma care intervention. Robust extraction of features from SCG signals in noisy environments is challenging due to low signal-to-noise ratios (SNR). In addition, lack of access to clean ground truth signals makes developing denoising algorithms even more difficult. In this work, we propose a novel deep learning-based approach for denoising SCG signals without requiring access to clean ground truth signals. Experimental results showed (1) enhancement in the SNR (approximately 10 dB and 9 dB increase for aortic opening (AO) and aortic closing (AC) regions of -10 dB SNR beats respectively), and (2) improvements in feature extraction accuracy (approximately 3x and 1.5x for AO and AC features of -10 dB SNR beats respectively) using the denoising model. Thus, the model effectively reduces noise, and improves the quality of SCG signals, leading to improved accuracy in feature extraction in noisy environments. This is a promising step forward in improving the quality and utility of SCG signals for clinical and research purposes.**

***Index Terms*—Seismocardiogram, Deep Learning, Noise Reduction, Signal to Noise Ratio**

I. Introduction

Hypovolemia, or low blood volume, due to trauma injuries can lead to hypovolemic shock, which continues to be the primary cause of death in both military and civilian populations [1], [2]. However, monitoring hypovolemia continues to be a difficult problem. Current monitoring solutions using traditional vital signs can be nonspecific and delayed in catching a patient's progression to shock [3]. Therefore, more robust solutions that can track blood volume loss accurately are necessary to enable early detection of hypovolemia and guided interventions to reduce the rate of mortality. Recent studies have demonstrated the efficacy of novel noninvasive systems and features for estimating blood volume loss [4], [5]. In particular, Zia et al. and Kimball et al. [2], [5] demonstrated that using key hemodynamic parameters from the electrocardiogram (ECG), photoplethysmogram (PPG), and seismocardiogram (SCG) allowed them to estimate blood volume loss by defining a measure called blood volume decompensation status (BVDS) that estimates the progression of blood loss to cardiovascular decompensation. Key to the accuracy of their blood volume loss model success was the inclusion of features derived from the SCG (pre-ejection period (PEP), left ventricular ejection time (LVET), and PEP/LVET), a noninvasive cardiovascular signal containing features correlated to aortic opening (AO) and aortic closing (AC) events. These features in particular are important for this task as they contain information about the left ventricular health [6].

This material is based in part on work supported by the Office of Naval Research under Grants N00014-21-1-2036 and N00014-22-1-2325

M. N., D. J. L., and O. T. I. are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (correspondence email: mnikbakht3@gatech.edu)

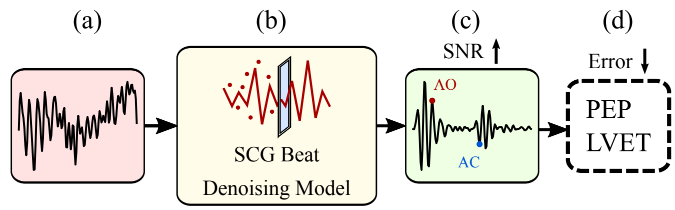

Fig. 1. Overview of the proposed SCG beat denoising model. (a) The noisy SCG beat with a low SNR is fed into the model as input. (b, c) The denoising model suppresses the noise artifacts while maintaining the signal power resulting in an output SCG beat with higher SNR. (d) This higher SNR SCG beat can be used for extracting PEP and LVET features more robustly with higher accuracy. SCG: Seismocardiogram; SNR: Signal to Noise Ratio.

However, the small amplitude of SCG signals makes them vulnerable to external vibrations and motion artifacts which can corrupt the AO and AC features. This susceptibility can result in unreliable estimation of SCG features, particularly in real-world scenarios such as when a patient is being transported to a hospital. As a consequence, health assessment algorithms that are primarily trained in controlled environments with lower noise levels may fail to function optimally in these noisier scenarios [7], [8]. As such, developing algorithms to mitigate the impact of these artifacts on SCG signals is essential to achieve robust and accurate SCG feature extraction in noisy environments. However, the unavailability of ground truth clean SCG signals under such conditions is a significant challenge in validating such denoising algorithms.

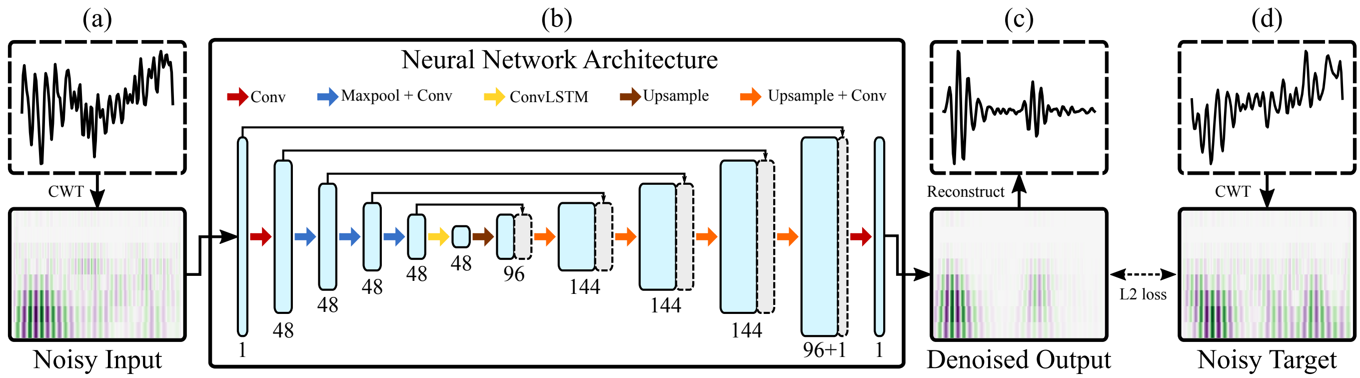

Fig. 2. The architecture of the proposed neural network. (a) The SCG beat waveforms are transformed to images using CWT. (b, c) The adapted Noise2Noise architecture proposed will receive the noisy SCG beats and output a higher SNR SCG beat as output. (d) During training, the target output is a noisy SCG beat with a different noise compared to the input noise and the model tries to minimize the loss between it's output and this target (c) resulting in a denoised output after convergence. CWT: Continuous Wavelet Transform; SNR: Signal to Noise Ratio; SCG: Seismocardiogram.

In this work, we introduce a deep learning architecture adapted from the Noise2Noise architecture, which was originally used for image restoration without clean labels [9]. They showed that learning to restore images from corrupted images by only using the corrupted examples (no image priors or corruption likelihoods) is possible [9]. Our proposed architecture is designed to restore clean SCG beats from a corrupted SCG beat, without requiring access to the corresponding clean ground truth labels. Fig. 1 shows an overview of the model proposed. The model takes a corrupted SCG beat as input (Fig. 1(a, b)), and is designed to preserve the cardiac information contained in the input while suppressing unwanted noise. The denoised beat output by the model (Fig. 1(c)) can then be used for extracting the AO and AC features, ensuring that the correct peaks and valleys are detected during feature extraction (Fig. 1(d)). This process can significantly improve the accuracy of SCG analysis and estimation health parameters in noisy or otherwise challenging environments.

The contributions of this work include: (1) Proposing a deep learning architecture and a novel methodology for restoring clean SCG beats from noisy SCG beats, without the need for clean ground truth SCG signals that can be particularly valuable in scenarios where obtaining clean ground truth data is challenging, which is often the case in real-world settings, and (2) demonstrating the effectiveness of our denoising model by validating its performance in improving the accuracy of AO and AC feature extraction from the SCG beats. Additionally, we show significant improvements in signal-to-noise ratio (SNR) after applying our denoising method, indicating the model's ability to effectively suppress unwanted noise and preserve essential cardiac information in the SCG signal.

II. Methods

A. Dataset

In this work, we used a dataset previously collected under a protocol approved by the Institutional Animal Care and Use Committees (IACUC) of the Georgia Institute of Technology (A100276), Translational Testing and Training Labs Inc. and the Department of the Navy Bureau of Medicine and Surgery [5]. This dataset contains recordings from an SCG sensor placed on the sternum and a 3-lead ECG sensor as well as catheter measurements from 6 anesthetized pigs going through blood draws of 7% at intervals to induce hypovolemia until 28% blood depletion or cardiac collapse [5]. Since collecting catheter measurements from human subjects for study purposes is difficult and the cardiovascular systems of humans and pigs are very similar, we chose this dataset for our study [5], [10]. In addition, SCG signals in this dataset contain a wide physiological range and since the pigs were anesthetized the SCG signals collected contain minimal motion noise, making it ideal for crafting our training data for the model proposed as explained in section II-B.

B. Pre-processing

In order to create the training dataset for the model proposed in this work (section II-C) we first filter the z-axis SCG signals using a bandpass Butterworth filter between 1-40 Hz and, using the R-peaks extracted from the ECG signals, we segment the SCG signals into cardiac cycles beats using each R-peak to 500 ms after as our bounds [7], [11]. After beat segmentation, since the beats are uncorrupted by motion artifacts, we add Gaussian noise filtered with the same bandpass filter to the SCG beats to generate noisy beats with -10 dB to 10 dB SNRs with respect to AO and AC regions power (since AO and AC complexes have different power levels) similar to the method in prior work [8], [9]. For each beat, we generated two noisy samples with two SNR values randomly picked from the -10 dB to 10 dB. Finally, we transformed each noisy 1D beat waveform pair to scalogram images (128x8) using the continuous wavelet transform (CWT) with 8 levels to create the input and label pair for training (Fig. 2(a, d)). In this work, we use the Morlet mother wavelet to generate the scalogram [12].

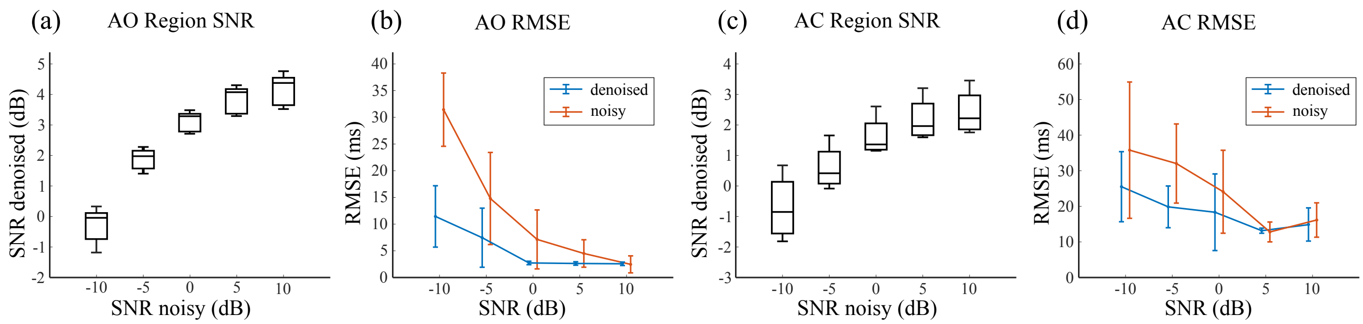

Fig. 3. (a, c) Box and whisker plots showing the mean SNR improvements in both the AO and AC regions of the SCG beats for all pigs (n=6) after denoising the beats using the denoising model proposed. (b, d) RMSE results on estimated PEP and LVET tracked features on all pigs (n = 6) for both noisy SCG beats and denoised SCG beats with different amounts of noise present in the noisy beats.

C. Model Architecture

Fig. 2(b) shows the architecture of the adapted U-Net based on the Noise2Noise approach [9], [13]. The model receives a noisy SCG input (Fig. 2(a)), and restores the clean SCG as output (Fig. 2(c)) trying to minimize the L2 loss (mean) between its output and the noisy target (Fig. 2(d)) [9]. We kept the activation functions the same (leaky ReLU with α = 0.1), and all convolutional layers have 'same' padding [9]. Our main alterations to the model are the channel sizes that we tuned during the hyperparameter tuning stage and, since the SCG signals have time features compared to images, an additional Convolutional LSTM (ConvLSTM) network [14], [15] to account for the temporal information in the SCG beats. To make sure the amplitude range of samples is in within [-0.5, 0.5], we scaled the CWT images by a constant factor.

We used an Nvidia GeForce GTX 1080 GPU for the training of the model. Adam optimizer was used with $\beta1$ = 0.9, $\beta2$ = 0.99, $\epsilon = 10e-8$, and learning rate of $3e-4$ [9]. A minibatch size of 4 was used and the weights were initialized with method introduced by He et al. [16]. Leave-one-subject-out (LOSO) cross validation (CV) was adopted for an unbiased estimate of the model performance. In this setup, one out of six subjects is held out and the model is trained on the beats of the remaining five subjects and used to denoise the beats of the held-out subject (total six models).

D. Validation

In order to assess the performance of our denoising model, we conducted two experiments. For both experiments, we created noisy beats by adding different levels of noise to the beats of the held-out subject. Specifically, we varied the SNR of the beats from -10 dB to 10 dB with 5 dB increments.

The first experiment focused on evaluating the performance of the model in enhancing the SCG beat's SNR. We fed each group of noisy beats for the held-out subject through the trained denoising model (trained on the remaining five subjects) and calculated the SNR after denoising for each group. To calculate the SNR, we subtracted the denoised beat from its clean ground truth and then calculated the power of this residual signal, which we refer to as the noise power [7]. We repeated this process for each held-out subject using its corresponding model (explained in section II-C). This allowed us to evaluate the effectiveness of the denoising model in reducing noise and improving the SNR of the SCG beats.

The second experiment involved evaluating the performance of the denoising model in improving the accuracy of AO and AC feature extraction from the SCG beats. Using a previously validated SCG feature detection algorithm based on consistent peak tracking [7], [17], we extracted the features from the denoised SCG beats and calculated the root-mean-squared error (RMSE) between extracted features and the ground truth AO and AC features that were previously validated for this dataset with the catheter gold standard measurements [5]. This process was repeated for all six subjects. By evaluating the accuracy of feature extraction before and after applying our denoising model, we can demonstrate the enhancement of feature quality achieved by our proposed approach.

III. Results

The results of the two experiments explained in section II-D are shown in Fig. 3. The box and whisker plots shown in Fig. 3(a, c) validate our first hypothesis that using the denoising model proposed we can enhance the SNR of SCG beats under different noise levels. We can observe that for -10 dB, -5 dB, and 0 dB SNR beat groups the model was able to enhance the SNR for both AO and AC regions. However, for 5 dB and 10 dB SNR beat groups, the denoised beat's SNR has decreased which can be due to the reason that the model hasn't seen clean ground truth labels during training. Note that for 5 dB and 10 dB groups, the SNR has dropped to approximately 4 dB and 2 dB for AO and AC portions respectively, which implies that the output beat still has a much higher signal power than noise. Fig. 3(b, d) agrees that the drop in denoised beat SNR doesn't affect the SCG features quality. We can also attribute the overall lower SNR values for AC region compared to AO region to the fact that, because the AC region has a lower power, it is usually more prone to noise [7], [11].

The RMSE results shown in Fig. 3(b, d) agree with our hypothesis that the proposed model increases the accuracy of feature extraction. We can see that for the AC features, the improvements are not as high as the AO feature accuracy. This difference is due to the fact that the AC features are usually lower in power compared to AO features and even minimal motion can corrupt them [7]. For AO detection accuracy we can observe approximately 3x, 2x, and 2x accuracy improvements for -10 dB, -5 dB, and 0 dB SNR groups respectively. For AC detection we can observe approximately 1.5x accuracy improvement for -10 dB, -5 dB, and 0 dB. For 5 dB and 10 dB in both AO and AC feature extraction, the improvements are not as significant due to the fact that the feature quality is already high even before denoising. However, we can see that for AO feature extraction, the accuracy improves slightly which means the feature quality is still enhanced proving that, although the denoising model reduces the SNR of these high SNR signals (Fig. 3(a, c)), the quality of the AO and AC features are not reduced (Fig. 3(b, d)).

These results suggest that the denoising model proposed enhances the reliability and accuracy of SCG-based feature extraction in the presence of noise. Such models are important for cardiovascular monitoring tasks, such as BVDS estimation, where accuracy and reliability are imperative and can be negatively affected by external vibration or motion artifacts commonly experienced during those tasks.

Some of the limitations of this work include: (1) this method is based on beat denoising therefore requiring beat segmentation steps and the ECG signal. (2) Only denoising of SCG signals corrupted by stationary noise was explored. Future work can adapt the architecture to explore the removal of non-stationary noise such as walking. (3) The noise was synthetically added to the signal to create the noisy pairs. Future work can explore ways to create naturally mixed signal and noise using systems that can replicate the SCG in noisy environments [18]. By employing such systems, it is possible to generate the same sequence of beats twice, which can then be naturally mixed with different noise samples to create the noisy input and target pairs with the same cardiac information but different noise.

Further, the architecture and approach proposed can be used to explore the effectiveness of such methods for other cardiomechanical signals such as ballistocardiogram (BCG) and phonocardiogram (PCG), or physiological waveforms such as ECG, PPG, and etc.

IV. Conclusion

In this work, we designed and validated a SCG beat denoising model trained without clean signals. Our results demonstrate that it is possible to restore a corrupted SCG beat without needing the corresponding clean SCG beat, which can be difficult to obtain. The model proposed can be used to reduce the noise in corrupted SCG beats while keeping the underlying cardiac content, resulting in enhanced SNR and improved accuracy of feature extraction. This would enable robust tracking of the SCG features in noisy environments for noninvasive patient health assessment helping healthcare professionals make more informed decisions and guide interventions.

References

[1] B. J. Eastridge, M. Hardin, J. Cantrell, L. Oetjen-Gerdes, T. Zubko, C. Mallak, C. E. Wade, J. Simmons, J. Mace, R. Mabry *et al.*, "Died of wounds on the battlefield: causation and implications for improving combat casualty care," *Journal of Trauma and Acute Care Surgery*, vol. 71, no. 1, pp. S4–S8, 2011.

[2] J. P. Kimball, J. S. Zia, S. An, C. Rolfes, J.-O. Hahn, M. N. Sawka, and O. T. Inan, "Unifying the estimation of blood volume decompensation status in a porcine model of relative and absolute hypovolemia via wearable sensing," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 9, pp. 3351–3360, 2021.

[3] S. R. Bruijns, H. R. Guly, O. Bouamra, F. Lecky, and W. A. Lee, "The value of traditional vital signs, shock index, and age-based markers in predicting trauma mortality," *Journal of Trauma and Acute Care Surgery*, vol. 74, no. 6, pp. 1432–1437, 2013.

[4] V. A. Convertino, G. Grudic, J. Mulligan, and S. Moulton, "Estimation of individual-specific progression to impending cardiovascular instability using arterial waveforms," *Journal of Applied Physiology*, vol. 115, no. 8, pp. 1196–1202, 2013.

[5] J. Zia, J. Kimball, C. Rolfes, J.-O. Hahn, and O. T. Inan, "Enabling the assessment of trauma-induced hemorrhage via smart wearable systems," *Science Advances*, vol. 6, no. 30, Jul. 2020.

[6] K. Sørensen, S. E. Schmidt, A. S. Jensen, P. Søgaard, and J. J. Struijk, "Definition of fiducial points in the normal seismocardiogram," *Scientific reports*, vol. 8, no. 1, p. 15455, 2018.

[7] D. J. Lin, J. P. Kimball, J. Zia, V. G. Ganti, and O. T. Inan, "Reducing the impact of external vibrations on fiducial point detection in seismocardiogram signals," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 1, pp. 176–185, 2021.

[8] D. J. Lin, A. H. Gazi, J. Kimball, M. Nikbakht, and O. T. Inan, "Real-time seismocardiogram feature extraction using adaptive gaussian mixture models," *IEEE Journal of Biomedical and Health Informatics*, 2023.

[9] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2noise: Learning image restoration without clean data," *arXiv preprint arXiv:1803.04189*, 2018.

[10] H. Schou, V. Perez de Sa, M. Sigurdardottir, R. Roscher, C. Jonmarker, and O. Werner, "Circulatory effects of hypoxia, acute normovolemic hemodilution, and their combination in anesthetized pigs," *The Journal of the American Society of Anesthesiologists*, vol. 84, no. 6, pp. 1443–1454, 1996.

[11] M. Nikbakht, A. H. Gazi, J. Zia, S. An, D. J. Lin, O. T. Inan, and R. Kamaleswaran, "Synthetic seismocardiogram generation using a transformer-based neural network," *Journal of the American Medical Informatics Association*, vol. 30, no. 7, pp. 1266–1273, 2023.

[12] A. Taebi and H. A. Mansy, "Time-frequency distribution of seismocardiographic signals: A comparative study," *Bioengineering*, vol. 4, no. 2, p. 32, 2017.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.

[14] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional lstm network: A machine learning approach for precipitation nowcasting," *Advances in neural information processing systems*, vol. 28, 2015.

[15] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026–1034.

[17] A. H. Gazi, S. Sundararaj, A. B. Harrison, N. Z. Gurel, M. T. Wittbrodt, M. Alkhalaf, M. Soudan, O. Levantsevych, A. Haffar, A. J. Shah *et al.*, "Transcutaneous cervical vagus nerve stimulation inhibits the reciprocal of the pulse transit time's responses to traumatic stress in posttraumatic stress disorder," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 1444–1447.

[18] M. Nikbakht, D. J. Lin, A. H. Gazi, and O. T. Inan, "A synthetic seismocardiogram and electrocardiogram generator phantom," in *2022 IEEE Sensors*. IEEE, 2022, pp. 1–4.